

IN THE SPECIFICATION

Please amend the paragraph beginning at page 8, line 6, as follows:

By way of a simple example, assume that a sentence decoded by the syllable speech recognizer 106 contains a sequence of syllables that are aligned to an audio (stored in unit 104) sentence whereby the sequence of syllables is represented as: S1, S2, S3, S4, S1, S4, S1, S2, S7, S8, S7. These syllables have been time stamped for times: t1-t2, t3-t4, t5-t6, tm-tn. Assume the audio sentence is represented as audio segments: aud1, aud2, aud3, aud4, aud5, aud6, aud7, aud8, aud9, aud10, aud11. Thus, the index data stored in unit 114 may be as follows: S1: aud1, aud5, aud 6⁷; S2: aud2, aud8; S3: aud3; S4: aud4, aud6; S7: aud9, aud11; S8: aud10. This means that a syllable S1 is stored in 1st, 5th and 6⁷th places (segments) in the audio sentence that is stored in the data storage 104. Therefore, to play segments that correspond to S1, one can go to corresponding locations in data storage 104 that are indicated in the index.